

**Improving Availability and Scalability  
in Clustered Application Servers**

Abstract of the Disclosure:

The invention relates to a technology of workload balancing for improved availability within a multitude of applications-servers and a multitude of application-clients interconnected with said application-servers by a communication network. A proposed method comprises a first-step, wherein an application-client is caching availability data of a subset of currently active application-servers as potential target application-servers. The method comprises a second-step, wherein for execution of an application-request the application-client selects an application-server from the subset based on a load-balancing decision of the application-client as target application-server. Finally the application-request is sent to the target application-server.